propagation across layers. We further propose an operational workflow and reporting standards applicable to multi-ancestry and multi-tissue settings. By reframing functional integration as an estimand-driven inference system rather than a collection of tools, this framework establishes a coherent theoretical foundation for bridging association discovery, causal probability, and mechanistic interpretation in complex trait genetics.

2 Integration of eQTL with Functional Phenotypes

In studies of complex traits, eQTL analyses provide an essential intermediate layer that links genetic variation to molecular phenotypes. Unlike GWAS, which captures statistical associations between loci and traits, eQTL focuses on how genetic variants influence gene expression or related molecular traits, thereby offering clues about potential regulatory mechanisms. In this sense, eQTL analyses do not by themselves establish causality, but instead help delineate the pathways through which genetic variants may act.

Within integrative frameworks, eQTL data are typically used in two ways. First, they provide molecular signals that can be compared with GWAS results through colocalization analyses. Second, they serve as a source of candidate instruments for downstream causal inference methods, such as Mendelian randomization. As such, eQTL analyses form a critical bridge between statistical association and functional interpretation.

2.1 cis-eQTL and trans-eQTL

eQTL are commonly categorized into cis- and trans-acting variants based on their genomic proximity to the target gene. Cis-eQTL are located near the gene they regulate and typically exert their effects through local regulatory elements such as promoters, enhancers, or untranslated regions. Large-scale datasets across multiple tissues have shown that cis-regulatory effects are widespread and relatively reproducible, with a substantial proportion overlapping GWAS loci (Liu et al., 2019; Wainberg et al., 2019). For this reason, cis-eQTL are often prioritized in integrative analyses as plausible links between genetic variants and gene expression.

In practice, cis-eQTL play two major roles. They can be used in colocalization analyses to assess whether GWAS and expression signals are likely driven by the same underlying variant. They also tend to provide more stable instruments for Mendelian randomization, enabling the evaluation of relationships between gene expression and phenotypic traits. At the same time, interpretation requires caution. Linkage disequilibrium may cause signals to spread across neighboring variants, and allelic heterogeneity can complicate the assignment of effects to specific loci. It is therefore common to combine eQTL results with fine-mapping, allele-specific expression, and functional annotations to strengthen the evidence.

Trans-eQTL, in contrast, affect genes located at a distance and often operate through indirect regulatory mechanisms, such as transcription factors, microRNAs, or chromatin interactions (Kirsten et al., 2015). These effects are typically weaker and more sensitive to cellular composition, environmental influences, and population structure, which makes their detection and interpretation more challenging. Recent work suggests that some trans effects arise through hierarchical regulatory relationships, where a cis-regulated gene acts as an upstream driver influencing downstream targets (Kvamme et al., 2025). This observation has practical implications for analysis strategies. One approach is to first identify candidate cis-regulatory variants, then explore their downstream impact using network-based methods or mediation analyses. In some cases, stepwise Mendelian randomization can be applied to evaluate relationships across multiple layers (Figure 1).

In plant systems, these challenges are often amplified by extended linkage disequilibrium, structural variation, and polyploid genomes. Incorporating multi-parent populations (such as NAM or MAGIC) and pangenome references can help reduce misattribution and improve robustness.

Image caption: Cis-eQTLs typically influence gene expression through proximal regulatory elements, such as promoters or enhancers, and are therefore more readily aligned with GWAS signals in integrative analyses. In contrast, trans-eQTLs act on distal genes through indirect mechanisms involving transcription factors, miRNAs, or chromatin interactions, often reflecting multi-layer regulatory processes and increased sensitivity to cellular and